

501.41034X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): MORIWAKI, et al.

Serial No.: Not yet assigned

Filed: January 16, 2002

Title: PACKET COMMUNICATION SYSTEM

Group: Not yet assigned



#2

LETTER CLAIMING RIGHT OF PRIORITY

Honorable Commissioner of
Patents and Trademarks
Washington, D.C. 20231

January 16, 2002

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Patent Application No.(s) 2001-184227, filed June 19, 2001.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Carl I. Brundage
Registration No. 29,621

CIB/alb
Attachment
(703) 312-6600

31 01-477

WN
J1011 U.S. PRO
10/046304
01/16/02

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office

#3

出 願 年 月 日
Date of Application:

2001年 6月19日

出 願 番 号
Application Number:

特願2001-184227

CERTIFIED COPY OF
PRIORITY DOCUMENT

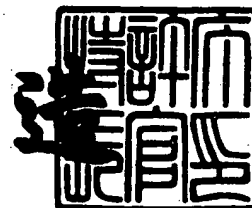
出 願 人
Applicant(s):

株式会社日立製作所

2001年10月19日

特 許 庁 長 官
Commissioner,
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3092562

【書類名】 特許願

【整理番号】 H01004771A

【あて先】 特許庁長官 殿

【国際特許分類】 H04L 12/56

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

 【氏名】 森脇 紀彦

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

 【氏名】 豊田 英弘

【特許出願人】

 【識別番号】 000005108

 【氏名又は名称】 株式会社 日立製作所

【代理人】

 【識別番号】 100075096

 【弁理士】

 【氏名又は名称】 作田 康夫

 【電話番号】 03-3212-1111

【手数料の表示】

 【予納台帳番号】 013088

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 パケット通信装置

【特許請求の範囲】

【請求項 1】

パケット通信装置であって、

第 1 の回線インタフェースと、

前記第 2 の回線インタフェースが収容する回線よりも低速な回線を収容する第 2 の回線インタフェースと、

クロスバススイッチと、

前記第 1 の回線インタフェース及び前記第 2 の回線インタフェースから、周期的に、パケット出力要求を受信し、その要求に基づき、前記第 1 の回線インタフェース及び前記第 2 の回線インタフェースに、前記クロスバススイッチに対するパケットの送出許可を送信するスケジューラとを有し、

前記第 1 の回線インタフェースと前記クロスバススイッチとの間のリンクの容量は、前記第 2 の回線インタフェースと前記クロスバススイッチとの間のリンクの容量よりも大きいパケット通信装置。

【請求項 2】

請求項 1 に記載のパケット通信装置であって、

前記第 1 の回線インタフェースと前記クロスバススイッチとの間のリンク数は、前記第 2 の回線インタフェースと前記クロスバススイッチとの間のリンク数よりも多いパケット通信装置。

【請求項 3】

請求項 1 又は請求項 2 の何れかに記載のパケット通信装置であって、

前記スケジューラは、同一周期において、前記第 1 の回線インタフェースから前記第 2 の回線インタフェースよりも多くのパケット出力要求を受信するパケット通信装置。

【請求項 4】

請求項 1 又は請求項 2 の何れかに記載のパケット通信装置であって、

前記スケジューラは、同一周期において、前記第 1 の回線インタフェースから

受信するパケット出力要求数の最大値と、前記第2の回線インタフェースから受信するパケット出力要求数と最大値との比は、前記第1の回線インタフェースと前記クロスバススイッチとの間のリンクの容量と、前記第2の回線インタフェースと前記クロスバススイッチとの間のリンクの容量との比と等しいパケット通信装置。

【請求項5】

請求項2に記載のパケット通信装置であって、

前記スケジューラは、同一周期において、前記第1の回線インタフェースから、最大で、前記第1の回線インタフェースと前記クロスバススイッチとの間のリンク数と等しい数のパケット送出要求を受信し、前記第2の回線インタフェースから、最大で、前記第2の回線インタフェースと前記クロスバススイッチとの間のリンク数と等しい数のパケット送出要求を受信するパケット通信装置。

【請求項6】

請求項1又は請求項2の何れかに記載のパケット通信装置であって、

前記スケジューラは、前記第1の回線インタフェースに前記第2の回線インタフェースよりも多くのパケットの送出許可を送信するパケット通信装置。

【請求項7】

請求項2に記載のパケット通信装置であって、

前記スケジューラは、前記第1の回線インタフェースに、最大で、前記第1の回線インタフェースと前記クロスバススイッチとの間のリンク数と等しいパケット数の送出許可を送信し、前記第2の回線インタフェースに、最大で、前記第2の回線インタフェースと前記クロスバススイッチとの間のリンク数と等しいパケット数の送出許可を送信するパケット通信装置。

【請求項8】

パケット通信装置であって、

複数の第1の回線インタフェースと、

それぞれが、前記第1の回線インタフェースの n 倍の速度を持つ複数の第2の回線インタフェースと、

前記複数の第1の回線インタフェースと前記複数の第2の回線インタフェース

と接続されるクロスバススイッチと、

前記複数の第 1 の回線インタフェース及び前記複数の第 2 の回線インタフェースから、周期的に、パケット出力要求を受信し、その要求に基づき、前記クロスバススイッチの制御を行うとともに、周期的に、前記複数の第 1 の回線インタフェース及び前記複数の第 2 の回線インタフェースに、前記クロスバススイッチに対するパケットの送出許可を送信するスケジューラとを有し、

前記複数の第 2 の回線インタフェースの各回線インタフェースは、第 1 の回線インタフェースと前記クロスバススイッチとを接続するリンク数の n 倍の数のリンクにより、前記クロスバススイッチと接続され、

前記スケジューラは、第 1 の回線インタフェースの入力側が 1 個の第 1 の回線インタフェースの出力側又は 1 個の前記第 2 の回線インタフェースの出力側と接続され、第 2 の回線インタフェースの入力側は最大 n 個の第 1 の回線インタフェースの出力側又は 1 個の第 2 の回線インタフェースの出力側と接続されるようにクロスバススイッチを制御するパケット通信装置。

【請求項 9】

請求項 8 に記載のパケット通信装置であって、

前記スケジューラは、同一周期において、最大で、第 1 の回線インタフェースに前記クロスバススイッチに対する送出を許可する最大のパケット数の n 倍の数のパケットの送出を第 2 の回線インタフェースに許可するパケット通信装置。

【請求項 10】

請求項 8 又は請求項 9 の何れかに記載のパケット通信装置であって、

前記複数の第 1 の回線インタフェースの各回線インタフェースの入力側及び出力側は、それぞれ前記クロスバススイッチと 1 本のリンクで接続され、前記複数の第 2 の回線インタフェースの各回線インタフェースの入力側及び出力側は、それぞれ前記クロスバススイッチと n 本のリンクで接続されるパケット通信装置。

【請求項 11】

請求項 10 に記載のパケット通信装置であって、

前記複数の第 1 の回線インタフェースの各回線インタフェース入力側は、前記 1 本のリンクと接続される 1 個の送信ドライバを有し、

前記複数の第 1 の回線インタフェースの各回線インタフェース出力側は、前記 1 本のリンクと接続される 1 個のレシーバを有し、

前記複数の第 2 の回線インタフェースの各回線インタフェース入力側は、前記 n 本のリンクと接続される n 個の送信ドライバを有し、

前記複数の第 2 の回線インタフェースの各回線インタフェース出力側は、前記 n 本のリンクと接続される n 個のレシーバを有するパケット通信装置。

【請求項 1 2】

パケット通信装置であって、

複数の第 1 の回線インタフェースと、

それぞれが、前記第 1 の回線インタフェースの n 倍の速度を持つ複数の第 2 の回線インタフェースと、

前記複数の第 1 の回線インタフェースと前記複数の第 2 の回線インタフェースと接続されるクロスバススイッチと、

前記複数の第 1 の回線インタフェース及び前記複数の第 2 の回線インタフェースから、周期的に、パケット出力要求を受信し、その要求に基づき、前記クロスバススイッチの制御を行うとともに、周期的に、前記複数の第 1 の回線インタフェース及び前記複数の第 2 の回線インタフェースに、前記クロスバススイッチに対するパケットの送出許可を送信するスケジューラとを有し、

前記複数の第 1 の回線インタフェースの各回線インタフェースは、速度 V のリンク速度で前記クロスバススイッチと接続され、

前記複数の第 2 の回線インタフェースの各回線インタフェースは、速度 $n \times V$ のリンク速度で前記クロスバススイッチと接続され、

前記スケジューラは、第 1 の回線インタフェースの入力側が 1 個の第 1 の回線インタフェースの出力側又は 1 個の前記第 2 の回線インタフェースの出力側と接続され、第 2 の回線インタフェースの入力側は最大 n 個の第 1 の回線インタフェースの出力側又は 1 個の第 2 の回線インタフェースの出力側と接続されるようにクロスバススイッチを制御するパケット通信装置。

【請求項 1 3】

請求項 1 2 に記載のパケット通信装置であって、

前記スケジューラは、同一周期において、最大で、第1の回線インタフェースに前記クロスバスイッチに対する送出を許可する最大の packets 数の n 倍の数の packets の送出を第2の回線インタフェースに許可する packets 通信装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、Internet Protocol (IP) などの可変長 packets や非同期転送モード（以下ATMという。）の固定長 packets （一般的にセルと呼ばれる）をスイッチングする packets データ通信装置に関する。

【0002】

【従来の技術】

近年、インターネットをはじめとするデータトラフィックは急激に増加している。また、従来専用線を使用して行なわれていたトランザクション処理など、高品質で、高信頼のサービスをインターネットで行おうとする動きも見られている。これに対応するため、伝送路だけでなく、packets データ通信装置の大容量化、高速化、高信頼化が必要とされる。

【0003】

packets データ通信装置のスイッチ方式としては大容量化の観点から入出力バッファ型スイッチが適していることが知られている。入出力バッファ型スイッチを用いる packets スwitchを開示するものとしては、"The Tiny Tera: A Packet Switch Core", Nick McKeown, Martin Izzard, Adisak Mekkittikul, William Ellersick, and Mark Horowitz, IEEE MICRO, January/February 1997（以下「文献1」という。）がある。文献1が開示するスイッチは、概ね図26に示すようなスイッチであると考えられる。 n 本の入出力ポートを有するクロスバスイッチ706の前段には、 n 個のポートカード701が設けられ、ポートカード701ごとに、入力バッファ703が配置される。入力回線700から入力された可変長 packets は固定長 packets （セル）単位に分割される。入力バッファ703にバッファリングされたセルは、スケジューラ705にて入出力ポート間の接続スケジューリングを行なった後に、各ポートカード701より出力されて、クロ

スバスイッチ706にてスイッチングされる。入出力ポート間の接続は1セル単位で変更される。特に、この構成においては、各入力バッファ703は出力方路別のキューバッファ (Virtual Output Queue (VOQ)) に分割されていて、スケジューラ705により出力指示された任意のキューバッファからの読み出しを可能とすることで、Head Of Line Blocking (HOL) によるスループットの低下を防止している。クロスバスイッチ706は、セル704を、例えば複数ビット単位にスライスして、複数のスイッチ面で並列処理する。

【0004】

また、従来のパケットデータ通信装置においては様々な種類の回線速度がサポート可能となっている。複数種類の回線をサポートする一般的なパケットデータ通信装置の主信号系構成を図3に示す。図3においては、クロスバスイッチ750は、2.4 Gbps 単位の複数の入力ポートと複数の出力ポートを有しており、入出力ポート間で $n \times n$ のスイッチングを行う。クロスバスイッチ750と回線インタフェースとの物理的な接続は、2.4 Gbps ドライバ (送信部) 730と、2.4 Gbps レシーバ (受信部) 731により行われる。本例においては、2.4 Gbps 回線インタフェース721のみならず、様々な種類の低速回線をサポートする例を示す。パケットデータ通信装置の回線収容効率を良くするため、回線インタフェースでは、低速回線を複数ポート分サポートするのが一般的である。図3においては、回線インタフェース722では600Mbps \times 4ポート、回線インタフェース723では150Mbps \times 16ポート、回線インタフェース724ではギガビットイーサ (1 Gbps) \times 2ポートを収容している例を示している。このように、低速回線については1つの回線インタフェース内に複数のポートを出来るだけ多く収容して、スイッチ可能なリソースを無駄にしないような構成を取る。

【0005】

【発明が解決しようとする課題】

今後、トラヒックの増加に伴い、さらなる高速回線をサポートする大容量スイッチが必要になると想定されるが、ネットワークのアクセス部分との接続、及び

従来装置との接続を考えると、従来の低速回線をもサポートする必要がある。

【0006】

図4に、スイッチ構成の一例を示す。クロスバススイッチ850は、40Gbps単位の複数の入力ポートと複数の出力ポートを有しており、入出力ポート間で $n \times n$ のスイッチングを行う。クロスバススイッチ850と回線インタフェースとの物理的な接続は、40Gbpsドライバ（送信部）830と、40Gbpsレシーバ（受信部）831により行われる。特に、数100Gbpsから数Tbpsクラスの大容量スイッチにおいては、クロスバススイッチ850と回線インタフェースとの物理的な接続が光インタコネクト等の光部品により行われる場合もある。図3と同様な考え方で、40Gbps回線インタフェース821のみならず、様々な種類の低速回線がサポートされる。クロスバススイッチ850のスイッチング単位としては40Gbpsであるが、例えば、2.4Gbps回線16本や、ギガビットイーサ回線を40本サポートすることは、搭載部品数および回線インタフェースの実装面積の都合上、分割損が生じて不可能となり、2.4Gbps回線8本（823）や、ギガビットイーサ回線を8本（824）程度に制限されて、回線インタフェースの容量密度が低くなる可能性がある。この場合、密度の低い回線インタフェースについては、クロスバススイッチ850との接続のために、40Gbps分のドライバ830、レシーバ831、もしくは光インタコネクトモジュールを搭載するのは冗長となるため、実装面積的にも、部品コストの観点からも望ましくない。この問題は、図4に例示したドライバ、レシーバの速度、収容する回線速度に依存するものではなく、高速回線と低速回線とを混在してスイッチに収容する場合に一般的に生じる問題である。

【0007】

また、文献1に示されているような従来の入出力バッファ型クロスバススイッチでは、入力ポートと出力ポートとの一対一接続を基本としている。つまり、複数の低速回線インタフェースと高速回線インタフェースが混在した状態で、これをクロスバススイッチで接続することを考えると、ある高速回線に入力されたトラヒックを複数の低速回線に同時に出力する、また、複数の低速回線に入力されたトラヒックをまとめて、一つの高速回線に同時に出力するような接続形態が不可能

であるため、スイッチの使用効率が著しく低下する場合がある。言い換えると、従来のクロスバスイッチにおいては、入出力ポート間の一对複数接続もしくは複数対一接続に対応していないため、低速回線インタフェースと高速回線インタフェースが混在した状態では、所望の出力ポートが空いているにもかかわらずデータを送出できない“ブロッキング”現象が発生する。

【 0 0 0 8 】

【課題を解決するための手段】

本発明の一側面は、各回線インタフェースとクロスバスイッチ間との接続ポート数を、各回線インタフェースに収容される回線の速度に応じて設定する。本発明の一実施例においては、低速回線インタフェース部と、クロスバスイッチとの入出力接続は1ポートを使用して接続し、高速回線インタフェース部と、クロスバスイッチとの入出力接続はnポートを使用して接続する。クロスバスイッチの入出力ポート間接続を決定するスケジューラは、低速回線インタフェースからは1つの送出要求リクエストを受信し、高速回線インタフェースからはポート毎に1個、つまりn個の送出要求リクエストを受信する。スケジューラは全てのポートから受信したリクエストをもとに入出力接続関係を決定し、ポート単位に送出可否を通知する。低速回線インタフェースでは、送出が認められた場合、通知された宛先出力に該当するパケットを送出する。高速回線インタフェースについては、n個のポートそれぞれにおいて、送出が認められた宛先出力に該当するパケットを送出する。ある高速回線インタフェースにおいて、複数の送出許可が1つの宛先出力に対して認められた場合には、宛先出力に該当するキューバッファから複数のパケットを連続して読み出しを行う。また、複数の送出許可が複数の宛先出力に対して認められた場合には、宛先出力に該当するキューバッファから、順次パケットの読み出しを行う。

【 0 0 0 9 】

また、本発明の他の側面は、スケジューラが、低速回線インタフェース部の入力側が1個の低速回線インタフェース部の出力側又は1個の高速回線インタフェース部の出力側と接続され、高速回線インタフェース部の入力側が最大n個の第1の回線インタフェースの出力側又は1個の第2の回線インタフェースの出力側

と接続されるようにクロスバススイッチを制御する。

【 0 0 1 0 】

【発明の実施の形態】

本発明によるパケット通信装置の実施例を説明する。

【 0 0 1 1 】

図 2 は、本発明のパケット通信装置の一実施例を示す。このパケット通信装置は、 $n \times n$ の交換を行うクロスバススイッチ 1 0、クロスバススイッチ 1 0 に接続される回線インタフェース 2 0 - 1 ~ 2 0 - n 、およびスケジューラ 4 0 とを有する。また、クロスバススイッチ 1 0、回線インタフェース 2 0、及びスケジューラ 4 0 は、制御部 5 0 により制御バス 5 1 を通じて、初期設定、統計情報収集、障害情報収集などの制御が行われる。

【 0 0 1 2 】

回線インタフェース 2 0 について説明する。回線インタフェース 2 0 の入力側は、入力処理部 2 1、VOQ (Virtual Output Queue) 2 3、およびVOQ制御部 2 2 を有する。回線インタフェース 2 0 の出力側は、VIQ (Virtual Input Queue) 3 3、VIQ制御部 3 1、および出力処理部 3 2 を有する。

【 0 0 1 3 】

図 1 4 は、入力処理部 2 1 の一実施例を示す。パケットデータは入力回線 4 0 を通じて装置に入力されると、光・電気信号変換部 (O/E) 2 1 - 1 にて、電気信号に変換される。その後、PHY 2 1 - 2 にてSONET (synchronous optical network) フレームなどの物理レイヤ処理が行なわれる。次に、L 2 処理部 2 1 - 3 にてパケットの抽出、エラーチェックなどのレイヤ 2 処理が行なわれる。その後、検索エンジン 2 1 - 4 にて、宛先 IP アドレスに基づき、出力ポート検索、品質クラス検索などの、レイヤ 3 処理が行なわれる。検索処理は、具体的には、検索エンジン 2 1 - 4 に接続された L 3 TABLE 2 1 - 5 を使用する。L 3 TABLE 2 1 - 5 には、予め宛先 IP アドレスと、出力ポート、品質クラス、次の転送先の IP アドレスであるネクストホップ IP アドレスとの対応関係がテーブル形式で格納されている。検索結果はパ

ケットのヘッダ部分に付与される。回線インタフェース20の出力側機能については後述する。

【0014】

以下では、クロスバスイッチに対して、40Gbps 1回線を収容する回線インタフェースカード（入出力40Gbps：以下40Gbps回線インタフェースと呼ぶ）2枚と、10Gbps×2回線（入出力20Gbps：以下20Gbps回線インタフェースと呼ぶ）を収容する回線インタフェース2枚が混在して搭載されている場合を例にして説明を行う。クロスバスイッチは搭載される4枚の回線インタフェースに対して入出力ポートの接続を行い、入力4×出力4のスイッチングを行う。なお、このパケット通信装置の回線スロットはスロットフリーを前提としており、40Gbps回線インタフェースと、20Gbps回線インタフェースのどちらも搭載できる構成とする。

【0015】

図1は、4×4スイッチングを行うパケット通信装置の一実施例を示す。このパケット通信装置は、4×4の交換を行うクロスバスイッチ10、クロスバスイッチ10に接続される回線インタフェース20-1～20-4、およびスケジューラ40とを有する。本実施例では、40Gbps×1ポートを収容する回線インタフェース20-1、20-3、および、10Gbps×2ポートを収容する回線インタフェース20-2、20-4が搭載されている。回線インタフェース20-1、20-3の入力側には20Gの送信ドライバ15がそれぞれ2個ずつ搭載されており、回線インタフェース20-1、20-3の出力側には20Gの受信ドライバ16がそれぞれ2個ずつ搭載されている。また、回線インタフェース20-2、20-4の入力側には20Gの送信ドライバ15がそれぞれ1個搭載されており、回線インタフェース20-2、20-4の出力側には20Gの受信ドライバ16がそれぞれ1個搭載されている。また、クロスバスイッチ10については回線インタフェーススロットに対して、2個の受信ドライバ16と送信ドライバ15がそれぞれ、スロットの受信側と送信側に配備されている。また、クロスバスイッチは、各入出力ドライバに対応して、回線インタフェーススロット当たり2つの物理入出力ポート（例えば、回線インタフェース20-1に相当

する入力ポートはIP11, IP12, 出力ポートはOP11, OP12)を持つ。40Gbps回線インタフェース20-1, 20-3とクロスバスイッチ10は2本の20Gbpsリンクで接続されており, また, 20Gbps回線インタフェース20-2, 20-4とクロスバスイッチ10は1本の20Gbpsリンクで接続されている。スケジューラ40は, 一定周期毎に制御線41を通じて全ての回線インタフェース20-1~20-4の packets 出力要求情報を収集して, これらをもとに回線インタフェース間の接続関係を決定し, 回線インタフェース20-1~20-4に packets の送出許可を出す。

【0016】

回線インタフェース間の接続関係を決定するためのスケジューラ40の一実施例について説明する。各出力回線インタフェースに対して, 複数の packets 出力要求が衝突した場合には, スケジュール40は, 巡回選択方式(ラウンドロビン)により接続すべき入力回線インタフェースを選択する。出力要求が優先度付けがされている場合には, 優先度の高い出力要求を優先するように入出力回線インタフェース間の接続関係が決定され, 優先度が重なった場合には, 同一優先度の出力要求に対して巡回選択が行われるようにすればよい。以上の選択を出力回線インタフェース毎に順次行うことで, 入力インタフェースと出力インタフェースとの接続関係が決定できる。

【0017】

図1では図示されていないが, 各回線インタフェースの入力側は, 図2で示したように, 入力処理部21, VOQ制御部22, 及びVOQ23を有しており, その出力側は, VIQ制御部31, 出力処理部32, 及びVIQ33を有している。

【0018】

VOQ制御部22の詳細, およびVOQ制御部とスケジューラ40との間の送受信フォーマットについて図5に示す。VOQ制御部22は, ヘッダ解析部221, バッファ管理部222, 及び送出要求作成部223とを有する。ヘッダ解析部221では, 入力された packets のヘッダを解析して, これをバッファ管理部222に通知する。バッファ解析部ではヘッダ解析部221から受け取った情報

を元に、パケットをVOQの所望のキューバッファ（出力回線インタフェースに対応）にライトするようにWA（Write Address）をVOQ 23に送出する。また、バッファ管理部222ではVOQ 23のパケット格納情報を送出要求作成部223に通知して、送出要求をスケジューラ40に送出する。スケジューラ40は各回線インタフェーススロットに対して2つの送出要求401, 402を受信可能な構成としてある。送出要求401, 402それぞれには、送出要求有効ビット（V）411, 412が付与されており、本ビットにより要求の有効／無効を判定する。具体的には図16に示すように40Gbps回線インタフェースについては、送出要求有効ビット（V）411, 412の双方を有効（“1”）にし、また、20Gbps回線インタフェースについては、図17に示すように、送出要求有効ビット（V）411, 412の片方のみ（本例では411）を有効（“1”）に設定しておく。つまり、スケジューラ40は、40Gbps回線インタフェースからは送出要求を2つ受け取ることができる。なお、40Gbps回線インタフェースの送出要求作成部では、VOQ 23内で送出可能なパケットが2個以上蓄積されている出力方路のキューバッファについては、送出要求401, 402の双方共に“1”を表示し（本例では出力#1）、送出可能なパケットが1個のみ蓄積されている出力方路のキューバッファについては、送出要求401のみに“1”を表示する（本例では出力#4）。図1の例においては、スケジューラ40はクロスバスイッチ40の物理ポートに対応した最大6個の送出要求を受け取る。スケジューラ40はこれら6個の送出要求に対して、クロスバスイッチ10の対応する6個の入出力ポート間でスケジューリングを行い、接続関係を決定する。接続関係が決定するとスケジューラ40は、送出許可を各回線インタフェース20のVOQ制御部22に対して返送する。スケジューラ40は各回線インタフェーススロットに対して2つの送出許可501, 502を送信可能な構成としてある。送出許可501, 502それぞれには、送出許可有効ビット（V）511, 512が付与されており、本ビットにより送出許可の有効／無効を判定する。具体的には図18に示すように40Gbps回線インタフェースについては、送出許可有効ビット（V）511, 512の双方を有効にし、また、20Gbps回線インタフェースについては、図19に示すように

送出許可有効ビット (V) 5 1 1, 5 1 2 の片方のみ (本例では 5 1 1) を有効に設定しておく。つまり, スケジューラ 4 0 は, 4 0 G b p s 回線インタフェースに対しては送出許可を最大 2 つ (送出許可 5 0 1, 5 0 2 のそれぞれ 1 つずつ) 送信し, 2 0 G b p s 回線インタフェースに対しては送出許可を 1 つのみ送信することができる。VOQ 制御部 2 2 内のバッファ管理部 2 2 2 においては, 1 つもしくは 2 つの送出許可を受け取ると, これに対応した VOQ 2 3 から 1 つもしくは 2 つの packets の読み出しが行なわれるように RA (Read Address) を VOQ 2 3 に送出する。

【0019】

以上の構成において, 2 0 G b p s 回線インタフェースと, 4 0 G b p s 回線インタフェース間の接続方法として, (1) 1 つの 4 0 G b p s 回線インタフェース入力が 1 つの 4 0 G b p s 回線インタフェース出力に接続 (2) 2 つの 2 0 G b p s 回線インタフェース入力が 1 つの 4 0 G b p s 回線インタフェース出力に接続 (3) 1 つの 4 0 G b p s 回線インタフェース入力が 2 つの 2 0 G b p s 回線インタフェース出力に接続, の 3 つの例について順に説明する。

【0020】

まず, 4 0 G b p s 回線インタフェース 2 0 - 1 の入力が 4 0 G b p s 回線インタフェース 2 0 - 3 の出力に接続される場合の回線インタフェース 2 0 - 1 の動作について図 6 を用いて説明する。スケジューラ 4 0 から制御線 4 1 を通じて, 出力側回線インタフェース 2 0 - 3 に対しての出力許可を 2 個受け取ると, VOQ 制御部 2 2 は VOQ 2 3 に RA を 2 つ与えて, 回線インタフェース 2 0 - 3 に相当するキューバッファ 2 3 3 から packets を 2 つ (A 1, A 2) シーケンシャルに読み出す。論理的には出力回線インタフェース 2 0 - 1 ~ 2 0 - 4 にそれぞれ対応したキューバッファ 2 3 1 ~ 2 3 4 からキューバッファ 2 3 3 をセレクト 2 3 0 にて選択して packets を読み出すことに相当する。読み出された 2 つの packets (A 1, A 2) は分離部 (DMX) 1 5 0 にて時間分離されて, それぞれ 2 0 G ドライバ (2 0 G - D R V) 1 5 を通じてクロスバスイッチの対応する入力ポート IP 1 1, IP 1 2 に対して出力される。これら 2 つの packets (A 1, A 2) はクロスバスイッチ 1 0 にて所望の出力ポート OP 3 1, OP 3 2 を

通じて、回線インタフェース 20-3 に対して出力される。回線インタフェース 20-3 出力側の動作について図 7 に示す。2 つのパケット (A1, A2) は、それぞれ 20G レシーバ (20G-DRV) 16 を通じて回線インタフェース 20-3 に入力される。その後、多重部 (MUX) 160 にて時間多重され、VIQ 内の該当する入力インタフェースに相当するキューバッファ (この場合は回線インタフェース 20-1 から入力されたパケットであるため、これに対応するキューバッファ 331) に格納されるように、VIQ 制御部 33 により制御される。VIQ 33 で元の可変長パケットに組み立てられた後、パケットは読み出されて出力処理部 32 へ送出される。

【0021】

ここで、出力処理部 32 の構成を図 15 を用いて説明する。入力された可変長パケットは L2 処理部 32-3 にてレイヤ 2 の処理が行なわれる。例えば、出力回線がイーサネットの場合には、次の転送先の IP アドレスであるネクストホップ IP アドレスから、接続先ルータのレイヤ 2 アドレス (MAC アドレス) を検索して付与する処理を行なう。ネクストホップ IP アドレスと、接続先ルータのレイヤ 2 アドレスの対応は、L2 TABLE 32-5 に格納されている。レイヤ 2 処理の終了後、可変長パケットは PHY 32-2 にて、例えば S O N E T フレームマッピング処理が行なわれ、その後電気・光信号変換部 (E/O) 32-1 にて、光信号に変換された後、出力回線 50 へ送出される。

【0022】

次に 2 つの 20Gbps 回線インタフェース 20-2 および 20-4 の入力が入力 1 つの 40Gbps 回線インタフェース 20-1 の出力に接続される場合の回線インタフェース 20-2 の動作について図 8 を用いて説明する。スケジューラ 40 から制御線 41 を通じて、出力側回線インタフェース 20-1 に対しての出力許可を 1 個受け取ると、VOQ 制御部 22 は VOQ 23 に RA を 1 つ与えて、回線インタフェース 20-1 に相当するキューバッファ 231 からパケット (B1) を 1 つ読み出す。論理的にはキューバッファ 231 ~ 234 からキューバッファ 231 をセレクタ 230 にて選択してパケット (B1) を読み出すことに相当する。読み出されたパケット (B1) は 20G ドライバ (20G-DRV) 15

を通じてクロスバススイッチ10の対応する入力ポートIP21に対して出力される。ここで、クロスバススイッチ10の入力ポートIP22に対しては何も接続されていない。パケット(B1)はクロスバススイッチ10にて所望の出力ポートOP11を通じて、回線インタフェース20-1に対して出力される。また、回線インタフェース20-4の動作について図9に示すが、これは図8で説明したのと同様な動作を行うので詳細説明を省略する。回線インタフェース20-4のキューバッファ231から読み出されたパケット(C1)はクロスバススイッチ10にて所望の出力ポートOP12を通じて、回線インタフェース20-1に対して出力される。回線インタフェース20-1出力側の動作について図10に示す。2つのパケット(B1, C1)は、それぞれ20Gレシーバ(20G-DRV)16を通じて回線インタフェース20-1に入力される。その後、多重部(MUX)160にて時間多重され、VIQ内の該当する入力インタフェースに相当するキューバッファ(パケットB1はキューバッファ332, パケットC1はキューバッファ334)に格納されるように、VIQ制御部33より制御される。最後に、40Gbps回線インタフェース20-3の入力が20Gbps回線インタフェース20-2および20-4の出力に同時に接続される場合の回線インタフェース20-3の動作について図11を用いて説明する。スケジューラ40から制御線41を通じて、出力側回線インタフェース20-2および20-4に対しての出力許可を受け取ると、VOQ制御部22はVOQ23にRAを2つ与えて、回線インタフェース20-2に相当するキューバッファ232からパケット(D1)を、また回線インタフェース20-4に相当するキューバッファ234からパケット(E1)をシーケンシャルに読み出す。読み出された2つのパケット(D1, E1)は分離部(DMX)150にて時間分離されて、それぞれ20Gドライバ(20G-DRV)15を通じてクロスバススイッチの対応する入力ポートIP31, IP32に対して出力される。パケット(D1)はクロスバススイッチ10にて所望の出力ポートOP21を通じて、回線インタフェース20-2に対して出力される。また、パケット(E1)はクロスバススイッチ10にて所望の出力ポートOP41を通じて、回線インタフェース20-4に対して出力される。クロスバススイッチ10の出力ポートOP22, OP42については何も接

続されていない。回線インタフェース20-2出力側の動作について図12に示す。パケット(D1)は、それぞれ20Gレシーバ(20G-DRV)16を通じて回線インタフェース20-2に入力され、VIQ内の該当する入力インタフェースに相当するキューバッファ(この場合はキューバッファ333)に格納されるように、VIQ制御部33より制御される。VIQ33で元の可変長パケットに組み立てられた後、パケットは読み出されて出力処理部32へ送出される。図13に回線インタフェース20-4出力側についての処理を示すが、これは図12と同様であるため説明は省略する。

【0023】

本発明を適用した回線インタフェースの実装イメージを次に示す。図24は40Gbps回線インタフェース70の場合である。クロスバススイッチ10と接続するための20GドライバLSI91および20GレシーバLSI92がそれぞれ2個ずつ搭載されており、これらの入出力信号はコネクタ93を通してバックパネル95に接続される。図25は20Gbps回線インタフェース80の場合である。クロスバススイッチ10と接続するための20GドライバLSI91および20GレシーバLSI92がそれぞれ1個ずつ搭載されており、これらの入出力信号はコネクタ93を通してバックパネル95に接続される。本例では、回線インタフェースとクロスバススイッチをバックパネルで接続する例を示したが、特に大容量パケット交換システムでは回線インタフェースとクロスバススイッチ間を電気ではなく光部品を使用して接続する場合も想定される。

【0024】

以上に示したように、本実施例によれば、大容量のパケット通信装置を構成する場合において、高速回線を収容した密度の高い回線インタフェースと低速回線を複数収容した密度の低い回線インタフェースが混在する場合、低速回線インタフェースには少ないドライバ・レシーバ(光接続の場合には光部品)を搭載すれば良く、高速回線インタフェースには、多くのドライバ・レシーバ(光接続の場合には光部品)を搭載することが実現可能であるようなコストリニアリティを有する回線インタフェースの実現が可能になる。

【0025】

また、複数の低速回線インタフェースと1つの高速回線インタフェース間でのスイッチの入出力ポート間の1対複数接続もしくは複数対1接続が可能となるような、パケット通信装置が提供可能になる。

【 0 0 2 6 】

以上説明した実施例では、高速回線インタフェースとクロスバススイッチとの間に2本のリンクを張り、低速回線インタフェースとクロスバススイッチとの間に1本のリンクを張る構成を示したが、高速／低速回線インタフェースに収容される回線の速度に応じて、高速回線インタフェースとクロスバススイッチとの間に n_H 本のリンクを張り、低速回線インタフェースとクロスバススイッチとの間に n_L 本のリンクを張る構成に拡張することも可能である。

【 0 0 2 7 】

他の実施例として、クロスバススイッチと回線インタフェース間の接続において、回線インタフェースの種別に応じて接続リンクの物理速度を変化させる方法について説明する。まず、回線インタフェースからクロスバススイッチ方向へのパケットの流れについて図20を用いて説明する。40Gbps回線インタフェース200-1には40Gbpsドライバ150が1つ搭載されており、20Gbps回線インタフェース200-2には20Gbpsドライバ15が1つ搭載されている。また、クロスバススイッチ100の受信側にはスロット単位に40Gbpsレシーバ160が1つ搭載されている。クロスバススイッチ100の受信側に搭載される40Gbpsレシーバ160では、40Gbpsのデータおよび20Gbpsのデータのどちらのデータが到着してもデータを受信することが可能な構成としておく。具体的には、20Gbpsデータ受信時には、40Gbpsデータ受信の半分のレートでデータの取り出し処理を行えば良い。クロスバススイッチ100のインタフェーススロットに対して、40Gbps回線インタフェース200-1が搭載される場合には、40Gbpsのデータ(A1, A2)を受信すると、分離回路(DMX)110にてこれを時間分離分離して、A1, A2をそれぞれクロスバススイッチ100内の入力ポートIP11およびIP12に入力する。これに対して、20Gbps回線インタフェース200-2が搭載される場合には、20Gbpsのデータ(C1)を受信すると、遅延回路(DLY)12

0にてこれをDMX110の処理時間分だけ遅延させて、C1をクロスバスイッチ100内の入力ポートIP21に入力する。入力ポートIP22には何も入力しない。これらの回線インタフェース速度に対してのモード切替は図22に示す構成で実現される。クロスバスイッチ100のスロット毎に、DMX110とDLY120の切替セレクタSEL1(101)、SEL2(102)を用意する。具体的には40Gbpsのデータが入力される場合には、データはDMX110にて時間分離されて、入力ポートIP11およびIP12へ入力される。また、20Gbpsのデータが入力される場合には、これがDLY140にて遅延されて、入力ポートIP11のみへ入力される構成とする。なお、回線インタフェースの搭載時に、回線インタフェースボードの容量に応じて異なるレベル線信号105(例えば、40Gbpsは、“1”, 20Gbpsは、“0”)が出力され、これによりSEL1(101)、SEL2(102)が設定される。また、SEL1(101)、SEL2(102)の設定はソフトウェアから行う構成にしても良い。回線インタフェース200間のデータ送出要求・送出許可信号のやりとり、スケジューリング方法については、前述の実施例と同様であるため、説明を省略する。

【0028】

次に、クロスバスイッチから回線インタフェース方向へのパケットの流れについて図21を用いて説明する。クロスバスイッチ100の送信側には回線インタフェーススロット単位に40Gbpsドライバ150が1つ搭載されている。また、これと接続される40Gbps回線インタフェース200-1には40Gbpsレシーバ160が1つ搭載されており、20Gbps回線インタフェース200-2には20Gbpsレシーバ16が1つ搭載される。クロスバスイッチ100の送信側に搭載される40Gbpsドライバ150では、40Gbpsのデータおよび20Gbpsのデータのどちらの速度のデータの送信を行うことも可能である。具体的には対向するレシーバ16が20Gbps対応の場合には、40Gbpsの半分のレートでデータの送出を行えば良い。クロスバスイッチ100に、40Gbps回線インタフェース200-1が搭載される場合には、クロスバスイッチ100の出力ポートOP11およびOP12から出力されるデータ

(A1, A2)を、多重回路(MUX)130にて時間多重し、40Gbpsの多重信号として回線側に送出する。また、20Gbps回線インタフェース200-2が搭載される場合には、出力ポートOP21からのデータ(C1)を受信すると、遅延回路(DLY)140にてこれをMUX130の処理時間分だけ遅延させて、C1を20Gbpsの信号として回線インタフェースに送出する。なお、出力ポートOP21からは何も出力しない。これらの回線インタフェース速度に対してのモード切替は図23に示す構成で実現される。クロスバスイッチ100の各スロット毎に、MUX130とDLY140の切替セレクタSEL3(103)を用意する。具体的には40Gbpsのデータが出力される場合には、出力ポートOP11およびOP12から出力されて、MUX130にて時間多重される。また、20Gbpsのデータが出力される場合には、これが入力ポートOP11から出力され、これがDLY140にて遅延される構成としておく。なお切替セレクタSEL3(103)については、回線インタフェースの搭載時に、回線インタフェースボードの容量に応じて異なるレベル線信号105(例えば、40Gbpsは“1”，20Gbpsは“0”)が出力され、これにより設定が実現される構成とする。また、SEL3(103)の設定はソフトウェアから行う構成としても良い。

【0029】

このように、本実施例によれば、大容量のパケット通信装置を構成する場合に、高速回線を収容した密度の高い回線インタフェースと低速回線を複数収容した密度の低い回線インタフェースが混在する場合、回線インタフェース速度に応じたドライバ・レシーバ(光接続の場合には光部品)を搭載すれば良く、コストリニアリティを有する回線インタフェースの実現が可能になる。また、複数の低速回線インタフェースと1つの高速回線インタフェース間でのスイッチの入出力ポート間の1対複数接続もしくは複数対1接続が可能な、パケット通信装置が提供できる。

【0030】

以上説明した実施例によれば、次のような効果が期待できる。

(1) 大容量のパケット通信装置を構成する場合に、高速回線を収容した密度の

高い回線インタフェースと低速回線を複数収容した密度の低い回線インタフェースが混在可能でかつ、回線インタフェース容量に比例した物理スイッチ間リンクを持つパケット通信装置を構成できる。

(2) 複数の低速回線インタフェースと1つの高速回線インタフェース間でのデータ転送が可能なパケット通信装置を提供できる。具体的には入出力回線インタフェース間の1対複数接続もしくは複数対1接続が可能となり、トラヒックのブロッキングが発生しないパケット通信装置を構成できる。

【0031】

【発明の効果】

高速回線を収容する回線インタフェースと低速回線を収容する回線インタフェースとが混在するパケット通信装置において、効率の良いパケットのスイッチングを行うことができる。

【図面の簡単な説明】

【図1】

本発明のパケット通信装置の機能ブロックを示すブロック図である。

【図2】

本発明のパケット通信装置の構成を示すブロック図である。

【図3】

従来のパケット通信装置の構成を示すブロック図である。

【図4】

従来のパケット通信装置の構成を示すブロック図である。

【図5】

本発明のパケット通信装置の回線インタフェースとスケジューラの構成を示すブロック図である。

【図6】

本発明のパケット通信装置の回線インタフェースカード（入力側）の構成を示すブロック図である。

【図7】

本発明のパケット通信装置の回線インタフェースカード（出力側）の構成を示

すブロック図である。

【図 8】

本発明の packets 通信装置の回線インタフェースカード（入力側）の構成を示すブロック図である。

【図 9】

本発明の packets 通信装置の回線インタフェースカード（出力側）の構成を示すブロック図である。

【図 1 0】

本発明の packets 通信装置の回線インタフェースカード（出力側）の構成を示すブロック図である。

【図 1 1】

本発明の packets 通信装置の回線インタフェースカード（入力側）の構成を示すブロック図である。

【図 1 2】

本発明の packets 通信装置の回線インタフェースカード（出力側）の構成を示すブロック図である。

【図 1 3】

本発明の packets 通信装置の回線インタフェースカード（出力側）の構成を示すブロック図である。

【図 1 4】

本発明の packets 通信装置の回線インタフェースカード（入力側）の構成を示すブロック図である。

【図 1 5】

本発明の packets 通信装置の回線インタフェースカード（出力側）の構成を示すブロックである。

【図 1 6】

本発明の packets 通信装置のスケジューリングで使用するデータフォーマット例である。

【図 1 7】

本発明のパケット通信装置のスケジューリングで使用するデータフォーマット例である。

【図 1 8】

本発明のパケット通信装置のスケジューリングで使用するデータフォーマット例である。

【図 1 9】

本発明のパケット通信装置のスケジューリングで使用するデータフォーマット例である。

【図 2 0】

本発明のパケット通信装置の回線インタフェースカードおよびクロスバスイッチの別構成を示すブロック図である。

【図 2 1】

本発明のパケット通信装置の回線インタフェースカードおよびクロスバスイッチの別構成を示すブロック図である。

【図 2 2】

本発明のパケット通信装置のクロスバスイッチの別構成を示すブロック図である。

【図 2 3】

本発明のパケット通信装置のクロスバスイッチの別構成を示すブロック図である。

【図 2 4】

本発明のパケット通信装置の回線インタフェースカードの一構成を示すブロック図である。

【図 2 5】

本発明のパケット通信装置の回線インタフェースカードの一構成を示すブロック図である。

【図 2 6】

従来の大容量パケットスイッチのブロック図である。

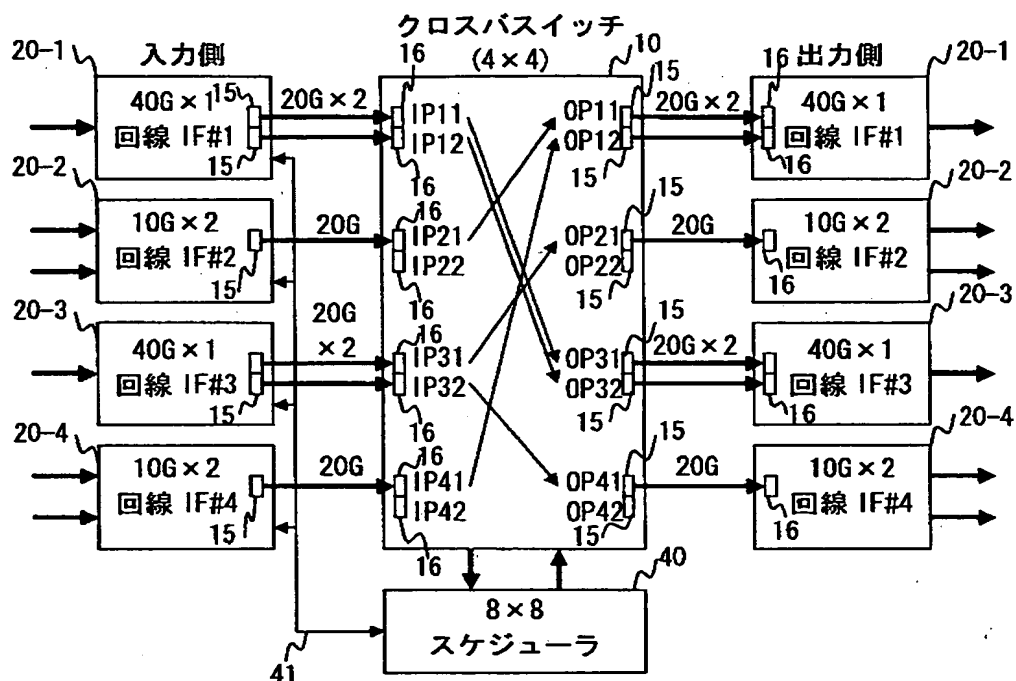
【符号の説明】

10...クロスバススイッチ, 20...回線インタフェース, 21...入力処理部
, 22...VOQ制御部, 23...VOQ, 31...VIQ制御部, 32...出力処
理部, 33...VIQ, 40...スケジューラ, 50...制御部

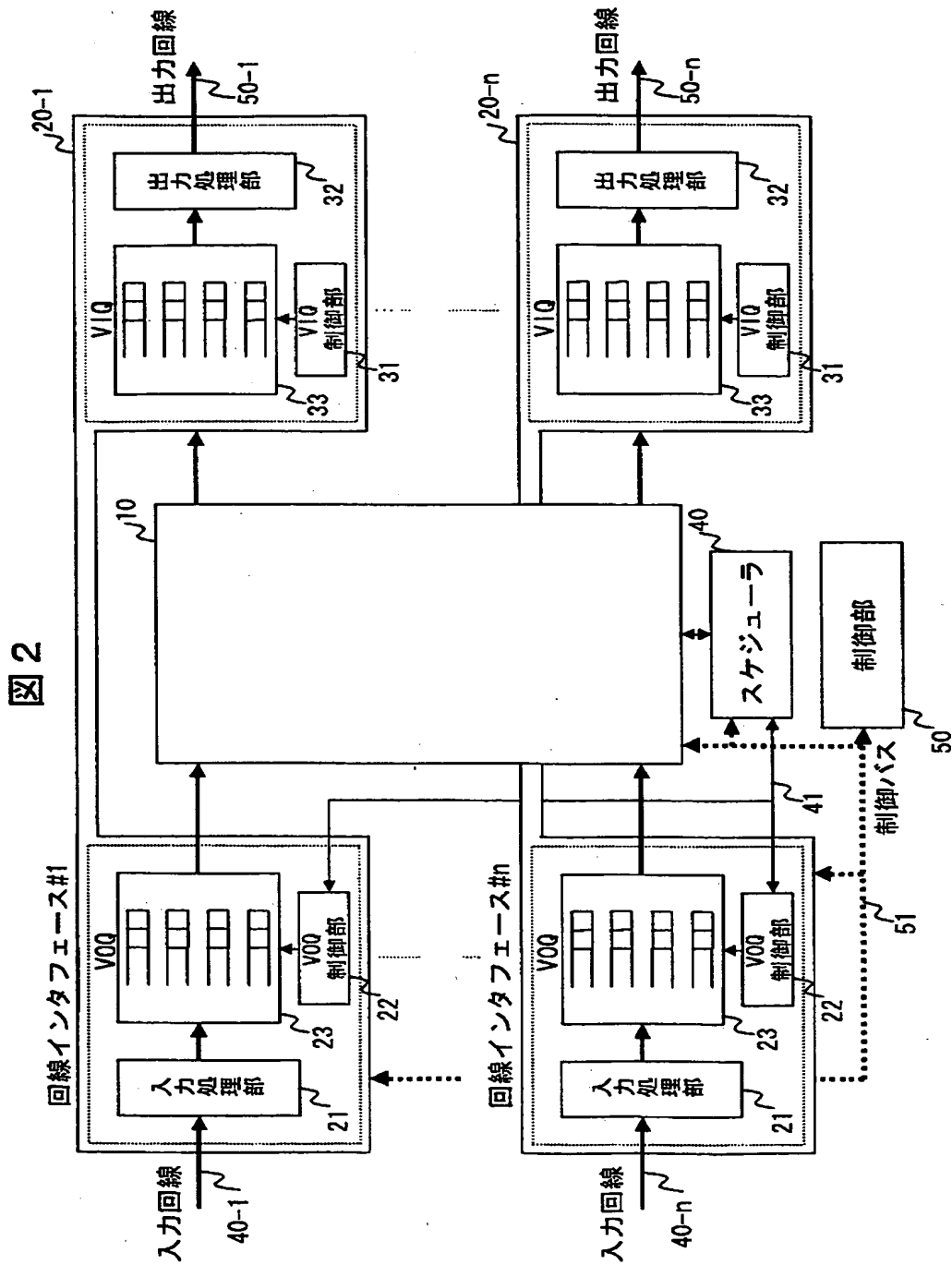
【書類名】 図面

【図 1】

図 1

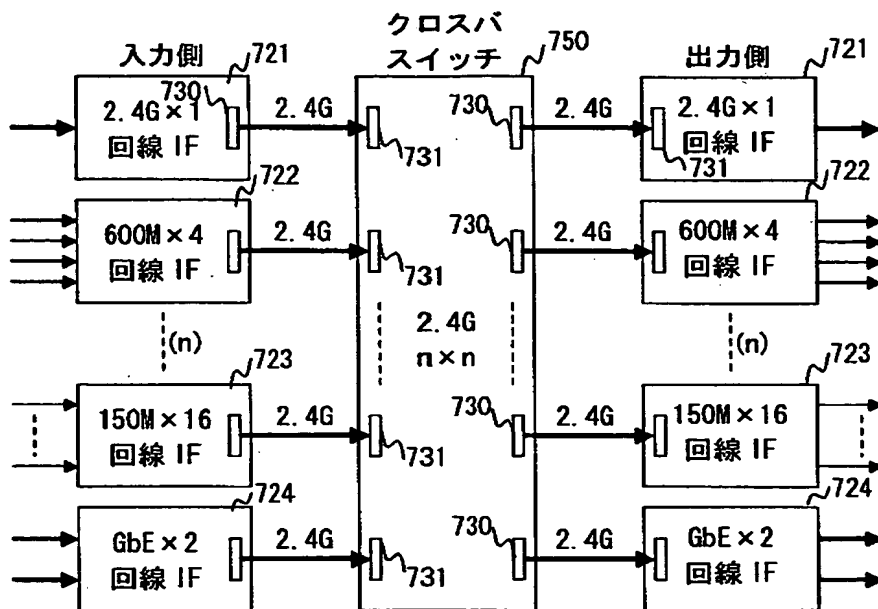


【図2】

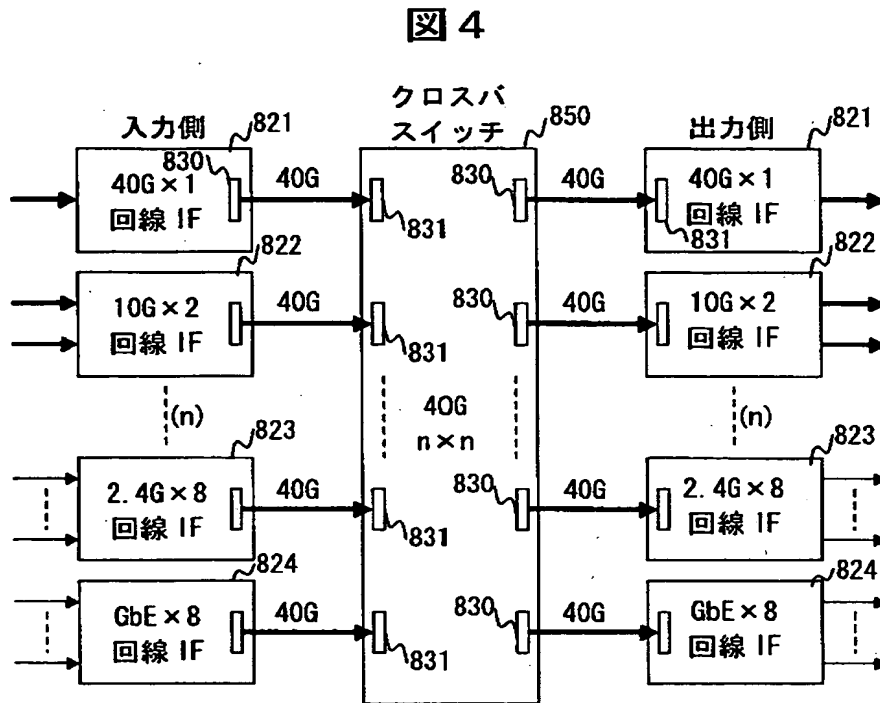


【図 3】

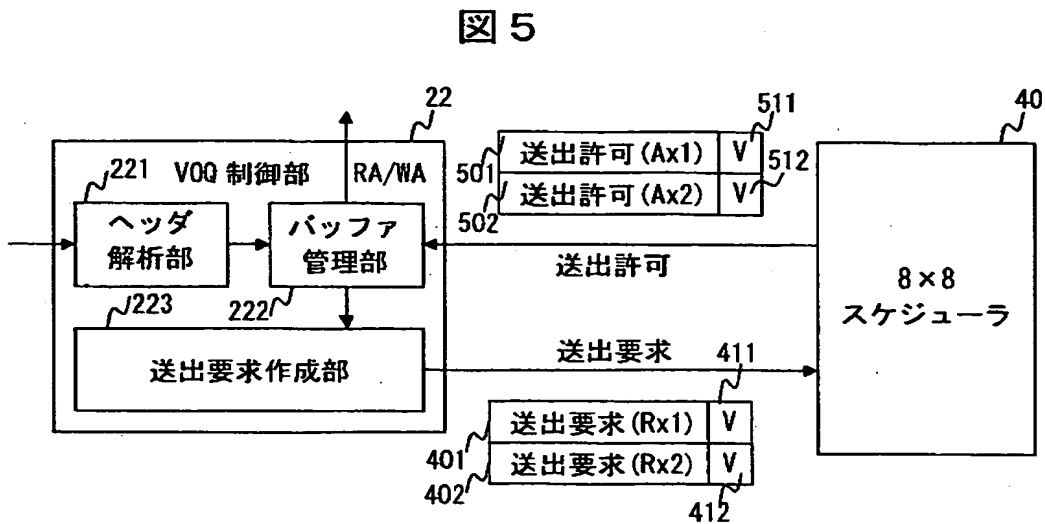
図 3



【図 4】

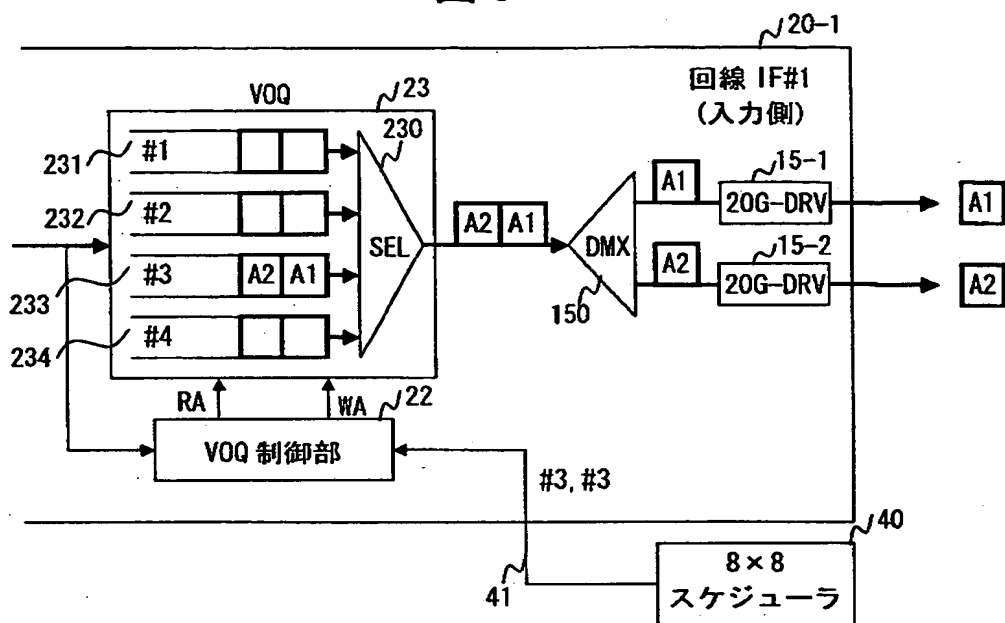


【図 5】



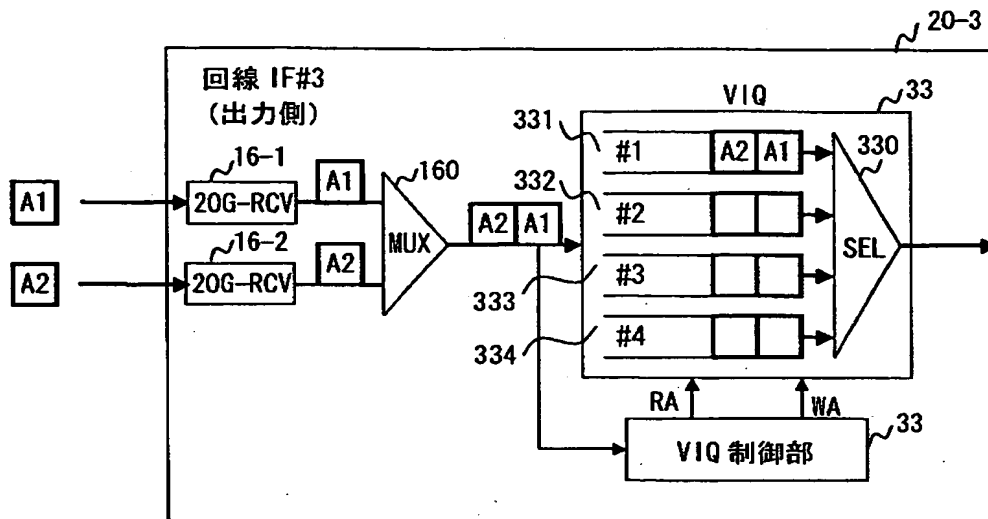
【図 6】

図 6



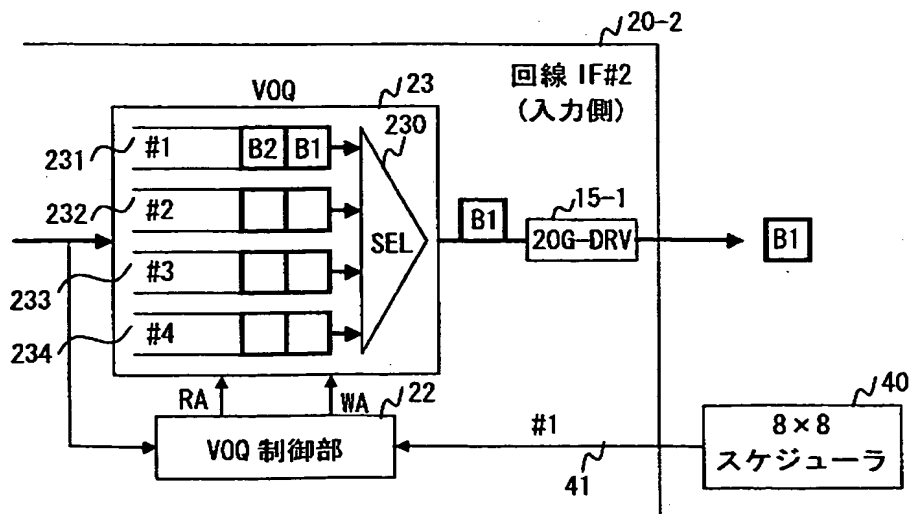
【図 7】

図 7

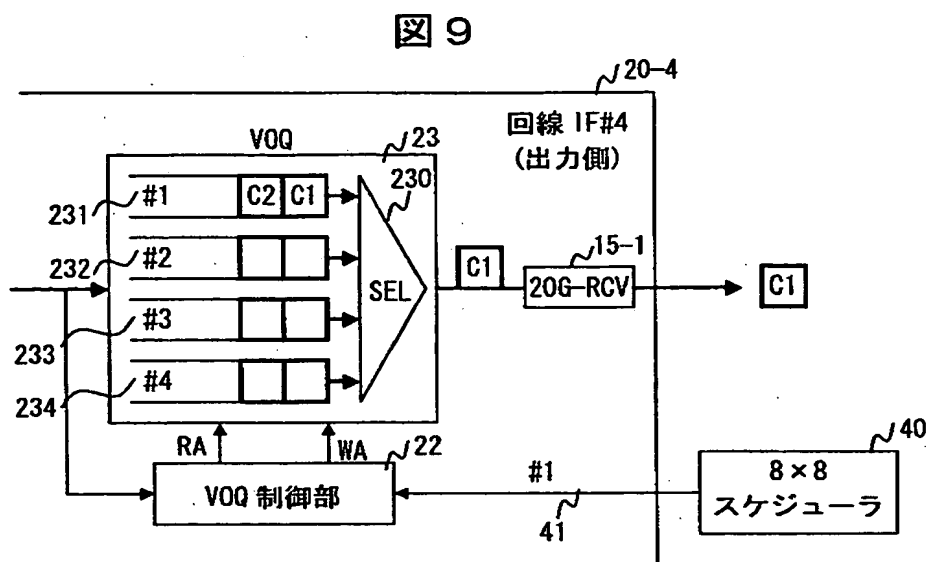


【図 8】

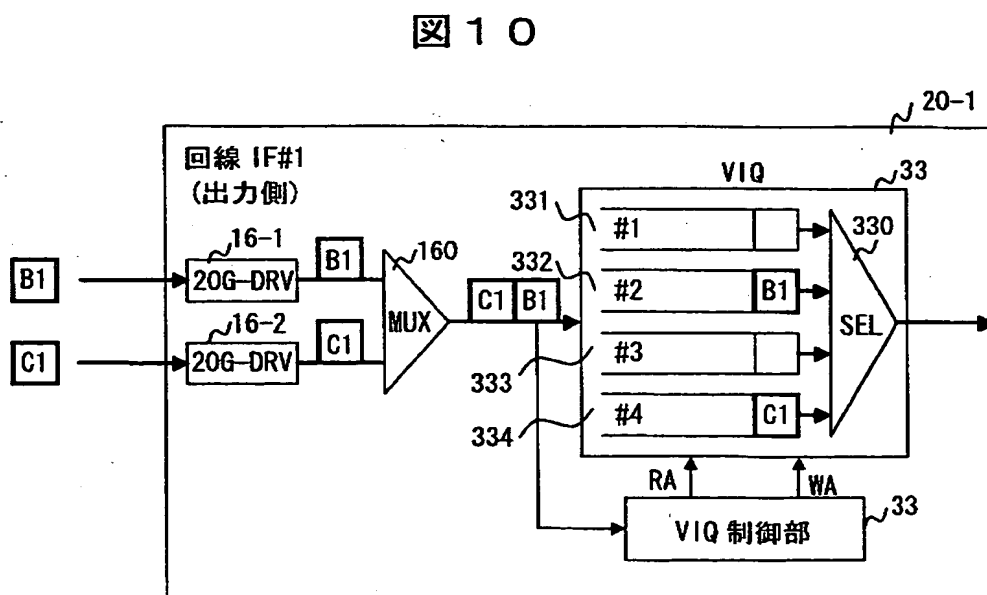
図 8



【図 9】

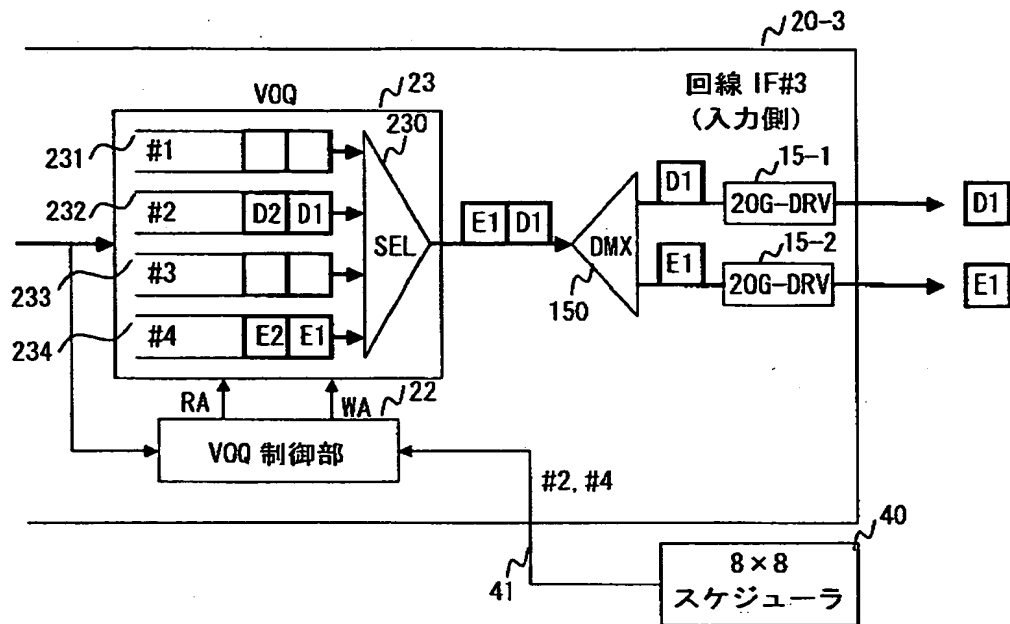


【図 1 0】



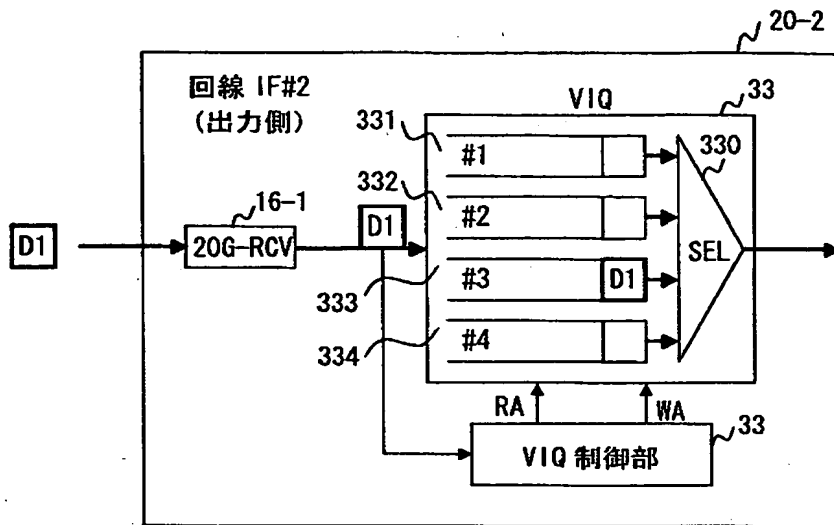
【図 1 1】

図 1 1



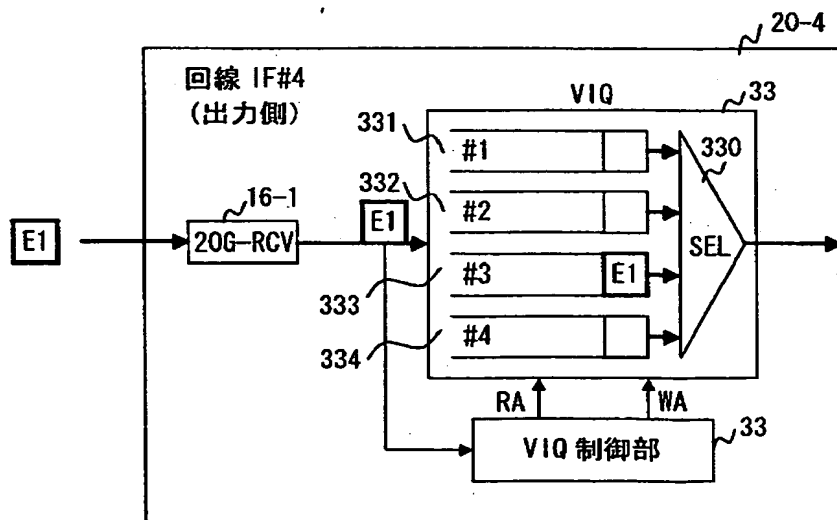
【図 1 2】

図 1 2



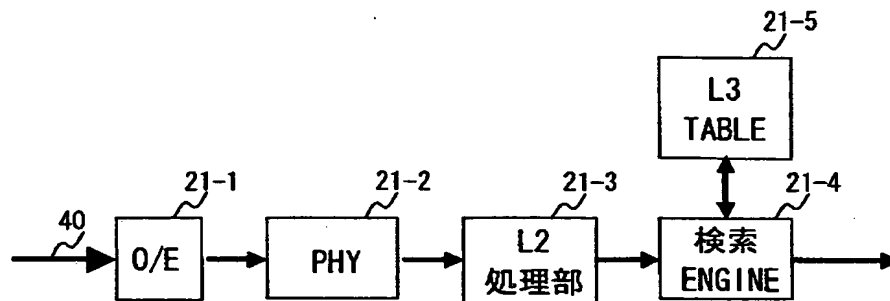
【図 1 3】

図 1 3



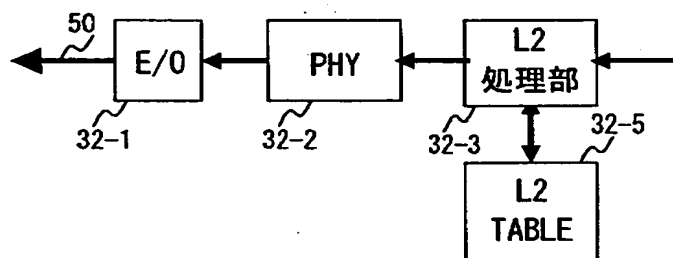
【図 1 4】

図 1 4



【図 1 5】

図 1 5



【図 1 6】

図 1 6

		出力#1	出力#2	出力#3	出力#4	
401	Rx1	1 (要求有)	0 (要求無)	0 (要求無)	1 (要求有)	1 (有効) 411
402	Rx2	1 (要求有)	0 (要求無)	0 (要求無)	0 (要求無)	1 (有効) 412

【図 1 7】

図 1 7

		出力#1	出力#2	出力#3	出力#4	
401	Rx1	1 (要求有)	0 (要求無)	1 (要求有)	0 (要求無)	1 (有効) 411
402	Rx2	d. c.	d. c.	d. c.	d. c.	0 (無効) 412

【図 1 8】

図 1 8

	出力#1	出力#2	出力#3	出力#4	
501 Ax1	0 (許可無)	1 (許可有)	0 (許可無)	0 (許可無)	1 (有効) 511
502 Ax2	0 (許可無)	0 (許可無)	0 (許可無)	1 (許可有)	1 (有効) 512

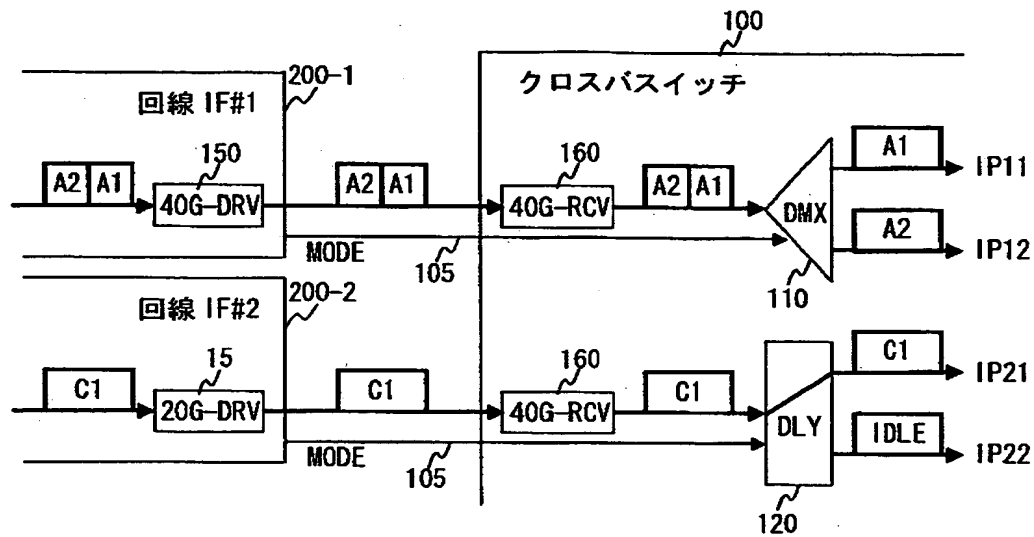
【図 1 9】

図 1 9

	出力#1	出力#2	出力#3	出力#4	
501 Ax1	1 (許可有)	0 (許可無)	0 (許可無)	0 (許可無)	1 (有効) 511
502 Ax2	d. c.	d. c.	d. c.	d. c.	0 (無効) 512

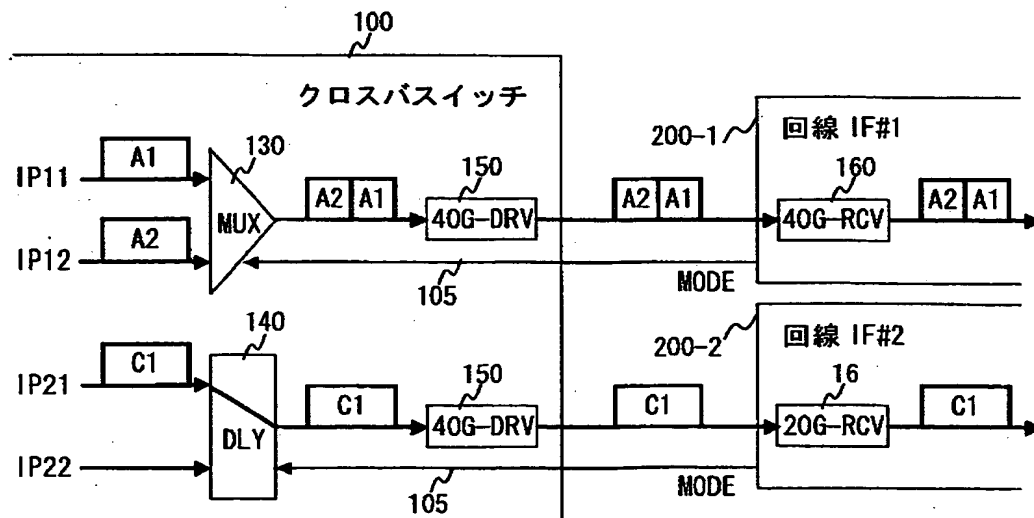
【図 20】

図 20



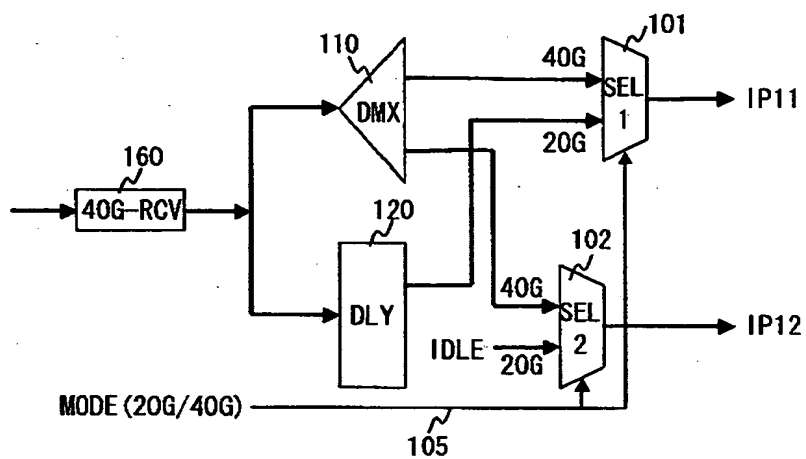
【図 21】

図 21



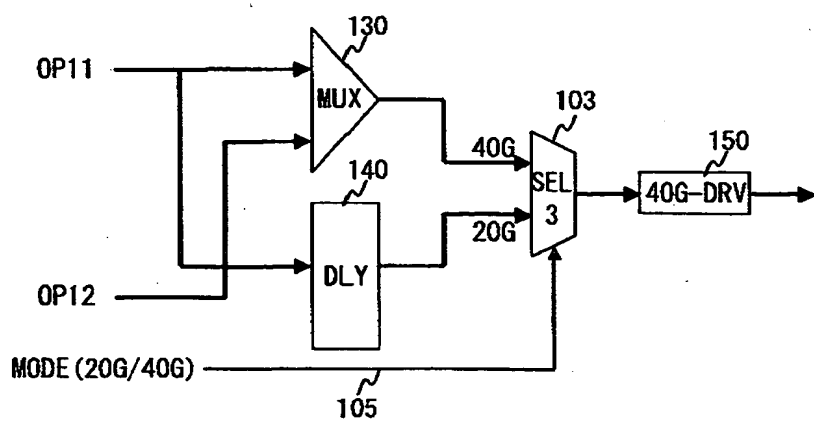
【図 2 2】

図 2 2



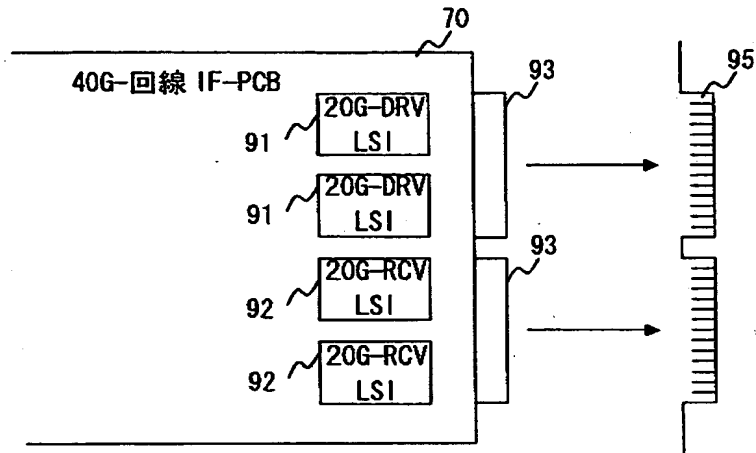
【図 2 3】

図 2 3



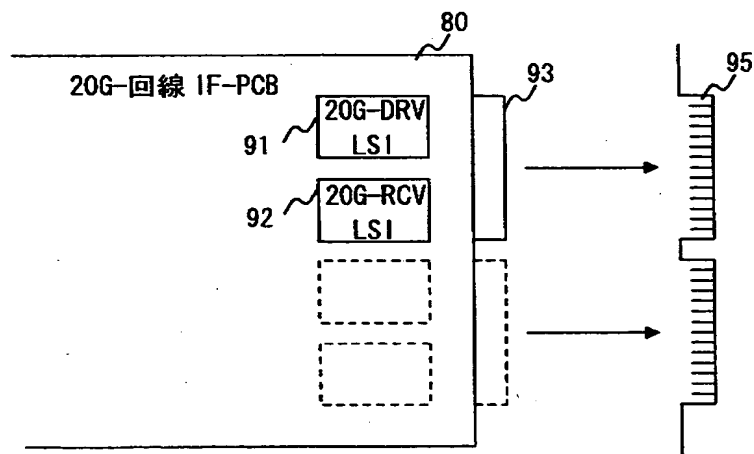
【図 24】

図 24



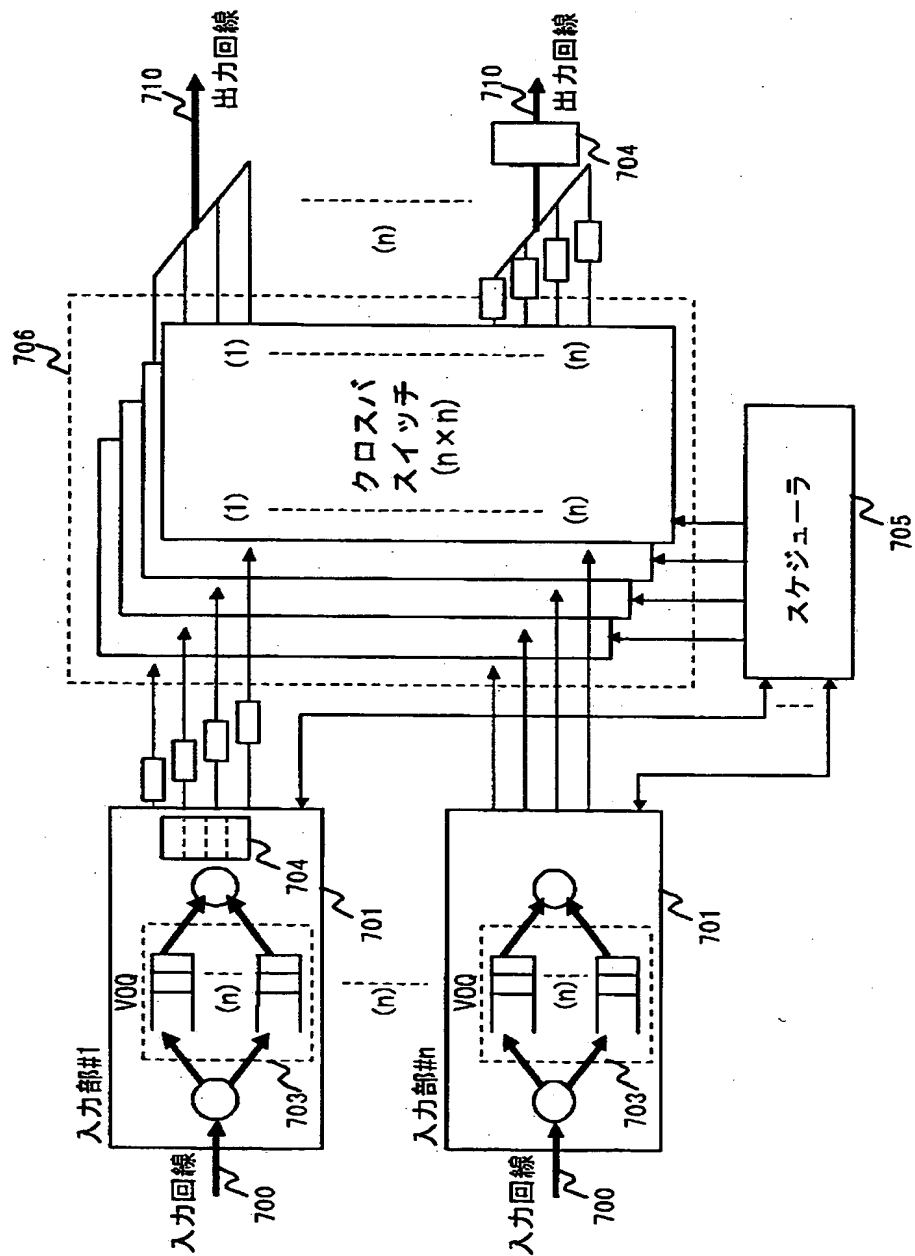
【図 25】

図 25



【図 26】

図 26



【書類名】 要約書

【要約】

【課題】異速度回線インタフェースを効率よく収容できるパケット通信装置を提供する。

【解決手段】本発明のパケット通信は、第1の回線インタフェースと、前記第1の回線インタフェースが収容する回線よりも低速な回線を収容する第2の回線インタフェースと、クロスバススイッチと、前記第1の回線インタフェース及び前記第2の回線インタフェースから、周期的に、パケット出力要求を受信し、その要求に基づき、前記第1の回線インタフェース及び前記第2の回線インタフェースに、前記クロスバススイッチに対するパケットの送出許可を送信するスケジューラとを有し、前記第1の回線インタフェースと前記クロスバススイッチとの間のリンクの容量を、前記第2の回線インタフェースと前記クロスバススイッチとの間のリンクの容量よりも大きくする。

【選択図】 図1

認定・付加情報

特許出願の番号	特願2001-184227
受付番号	50100880473
書類名	特許願
担当官	第八担当上席 0097
作成日	平成13年 6月20日

<認定情報・付加情報>

【提出日】	平成13年 6月19日
-------	-------------

出願人履歴情報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所